Google Earth Engine, an innovative technology for forest conservation

Thuan Sarzynski[1,2], Anushka Rege[3], Smita Bodhankar Warnekar[4], Sifeng Wu[3,5], Janice Ser Huay Lee[3]

[1] CIRAD, UMR DIADE, Montpellier, France.
[2] DIADE, Supagro Montpellier, CIRAD, IRD, Montpellier, France.
[3] Asian School of the Environment, Earth Observatory of Singapore, Nanyang Technological University of Singapore, Singapore.
[4] Independent researcher.
[5] Complexity Institute, Interdisciplinary Graduate School, Nanyang Technological University, Singapore.

Abstract

Millions of satellite images of the Earth's surface are produced each day. However, these images can be difficult to access and analyze as they are scattered across different websites, and processing them can take weeks or months due to the computing limits of office computers. Google Earth Engine (GEE) can help solve these problems. GEE is a web application where all open-access images of the Earth's surface are compiled. The application benefits from the computing power of thousands of computers at Google's data centers, thus reducing the time needed to conduct land cover analysis at national, regional and planetary scales. Since its creation in 2010, the application has been used in a wide range of fields such as agriculture, economics, medicine, and forestry. For instance, in 2013, Dr. Matthew C. Hansen (University of Maryland, the United States of America) has automated the analysis of Earth's forest cover using GEE, allowing analyses of forest cover loss to be constantly updated (Hansen et al. 2013).

During his eight-month master's thesis, the first author (TS) mapped oil palm plantations in Sumatra, Indonesia using GEE. The map had a high overall accuracy and compared well with the official Indonesian statistics. The code and the input data are freely available on TS's GitHub account for future experimenters to apply in different years or regions. Similarly, the co-authors (SBW, AR and JL) are currently using GEE to map cashew monocultures in Sindhudurg district (southern Maharashtra state, India). These studies, made possible by the use of GEE, were a first step toward the close monitoring of oil palm and cashew plantations, which are important drivers of deforestation in Indonesia and India.

Such land use monitoring could be programmed to be constantly self-updating like in Dr. Hansen's project. A small number of resource persons could be trained to use GEE and provide inputs for such programs.

These land use monitoring tools can help enforce forest protection laws and identify companies sourcing oil palm or cashew from protected areas. In addition, the information acquired using GEE can assist organizations managing oil palm sustainability certification schemes in excluding companies that drive deforestation or source material from protected areas. People equipped with GEE skills can play a key role in influencing landscape planning and making adequate policies in the forestry sector. These tools can help prevent or reduce CO2 emissions due to commodity-driven deforestation, as well as protect the abundant biodiversity present in natural forests.

Keywords: *remote sensing, google earth engine, deforestation, land-use change, oil palm.*

Introduction and background

Background

Deforestation for agricultural expansion is a persistent global threat to natural ecosystems and contributes to carbon dioxide emissions and climate change. The expansion of agricultural land has been driven by increasing human population and growing demand for food, fuel and fiber. The challenges associated with monitoring the conversion of natural ecosystems to agricultural land can be informed using improved technologies for land cover and land use mapping.

Google Earth Engine (GEE) was developed by Google in 2010 to help land planners acquire and analyze satellite imagery at a planetary scale. Millions of satellite images of the Earth's surface are produced each day. However, they are often scattered across different sources in various data formats and can be challenging and time-consuming to analyze due to the computing limitation of personal computers.

GEE is a cloud-based platform powered by Google Cloud Platform and can be operated seamlessly with other online services provided by Google. With the help of GEE, ready-to-use datasets can be easily loaded and processed. Then, the processed data and analyzed results can be stored and exported for future use. GEE has its own application programming interface (API), which includes a library of operators in both JavaScript and Python computer languages, making it accessible to a wider community. The platform increases accessibility to high-performance computing resources for processing large geospatial datasets (Gorelick et al. 2017). GEE provides users with access to the computational power of thousands of computers from Google's data centers through its online code platforms.

The GEE code platform[1] works with JavaScript, while Python users can deploy the computational power through Google Colab,[2] an online Python application for general purposes that also allows the use of GEE, as well as free access to the computing power of Google.

It is worth mentioning that among its vast number of libraries, GEE incorporates most of the state-of -the-art machine learning (ML) methods, which can be easily called and applied to the loaded or processed geospatial datasets as the algorithms have been packed into three packages (namely Classifier, Clusters and Reducer). For example, the widely-used classification methods for land use classification, like random forest classifier and other tree-based classifiers, are all included in the Classifier package and can be called with a few lines of code.

When dealing with more complex models or larger training datasets, GEE also works well with TensorFlow, which is an open access library to facilitate efficient ML analysis. The

1 Accessible at https://code.earthengine.google.com/

2 See: https://research.google.com/colaboratory/faq.html

built-in functions and packages make the learning curve of GEE much less steep. Moreover, the active online community of GEE can be greatly helpful to learners. Not only do users make use of previous code and packages, but GEE has also recently created a feature to integrate any code into a shareable user interface called Earth Engine Apps. Such apps can use open access code developed by users and leverage GEE's data catalog and analytical power to detect deforestation and land use change.

Since its launch, GEE has been used worldwide in various fields spanning agriculture, ecology, economics, forestry and medicine (Kumar and Mutanga 2018). GEE is a very promising innovation that, by providing large datasets and high computation power to users, can facilitate the mapping of Earth's land cover at national, regional and planetary scales (Hansen et al. 2013).

Objective of the article

This article aims to introduce readers to GEE with two examples of its uses for land cover mapping: (i) industrial oil palm detection in Indonesia; and (ii) differentiating cashew from forest cover in India. These two case studies provide insights about the implementation, impact and constraints of this innovative technology for monitoring agricultural land cover that commonly encroaches into forests.

Description of the innovation

Industrial oil palm detection in Indonesia

The first author (TS) mapped oil palm plantations for the year 2015 in Riau, Jambi, and South Sumatra, three provinces of Indonesia known for their large areas dedicated to oil palm (Sarzynski et al. 2020).

Imagery from two optical satellites, Landsat 7 and Landsat 8, and one radar satellite, Palsar, were all sourced and pre-processed on GEE. The analysis included running the classification of land cover separately for each type of satellite imagery and when optical and radar imagery were combined (Joshi et al. 2016). Apart from the wavelength bands used in the satellite images, a range of indices (e.g. NDVI – normalized difference vegetation index,[3] EVI – enhanced vegetation index,[4] etc.) were calculated and included in the classification process, which involved a random forest (RF) algorithm,[5] a common ML method based on building a large number of decision trees to classify six land cover types: forest, forest-shrub mosaic, oil palm, bare-ground crop, built-up area, and water. Both the analysis and accuracy assessment were conducted using the GEE API.

The overall accuracy of the resulting map, which is the number of correctly classified pixels over the total number of pixels, reached 84 percent. The pixels used to calculate this overall accuracy were selected for each land cover type from polygons we delineated in Google Earth Pro. Our map estimation of oil palm area in the three provinces (4.4 million ha) compared well with the official Indonesian statistics (4 million ha). We compared our technique with those in two other studies to map oil palm in the same area. Our technique provided a closer estimate of oil palm extent to the official statistics than the other studies, but with a substantial number of discrepancies with the two maps of the other studies. Our map of oil palm was compared to a map of protected areas to detect potential illegal deforestation and encroachment. This comparison showed that oil palm plantations encroached on 3 percent of the protected areas in these provinces in 2015.

This work can contribute to closely monitoring oil palm plantations and better identifying whether and where oil palm plantations have expanded over forests and caused deforestation. The code and source data are freely available on GitHub online repository[6] for other researchers to apply in different years or regions. Subsequent maps drawn from different years can highlight changes in

3 See: https://en.wikipedia.org/wiki/Normalized_difference_vegetation_index

4 See: https://en.wikipedia.org/wiki/Enhanced_vegetation_index

5 See: https://en.wikipedia.org/wiki/Random_forest

6 See: https://github.com/thuansarzynski/GEE_CombinedLandsatSAR_oilpalm

land use and land cover over time and inform policymakers and law enforcement officials on where natural resources are threatened. For example, a sustainable certification program could exclude oil palm companies establishing plantations on forested land.

Differentiating cashew monocultures from forests in India

While many studies have already mapped crops such as oil palm and coffee, cashew monoculture has not yet received much research attention. Cashew is currently grown across 33 tropical countries that also host high levels of biodiversity (FAOSTAT).[1] The co-authors (SBW, AR and JL) are currently using GEE to map cashew monocultures in Sindhudurg district (southern Maharashtra, India). India is the second-largest cashew producer in the world, and Sindhudurg is one of the country's top cashew-producing districts (C-DAP Sindhudurg 2012; FAOSTAT[7]). The forests in Sindhudurg fall in the Western Ghats global biodiversity hotspot[8]. Most of them are privately-owned forests and serve as a wildlife corridor (Myers et al. 2000; Punjabi and Kulkarni 2015). In recent years, cashew monocultures have increasingly replaced forests in this landscape.

The authors obtained and pre-processed optical (LANDSAT-8) and radar (SENTINEL-1) imagery from February to May 2020 using GEE. Pre-processing involved speckle filtering and mosaicking the image to remove cloud cover. For analysis, random forest (RF) and classification and regression tree (CART)[9] algorithms were run on combined optical and radar imagery using polygons of the following land cover categories: forest, monoculture cashew, barren land, crop land, and water bodies (delineated in Google Earth Pro[10] using on-ground field observations).

The total extent of land under cashew cultivation in Sindhudurg according to our preliminary GEE analysis was 118,000 ha using the RF algorithm, and 74,500 ha using the CART algorithm. In 2018, Maharashtra had 191,450 ha of land under cashew cultivation, of which Sindhudurg is known to contribute approximately 41.1 percent or 78,685.9 ha (Sengar et al. 2012; Directorate of Cashew nut and Cocoa Development 2018).

The 58 percent difference between RF and CART algorithms is due to differences in how they work, with the CART algorithm being more robust for outlier predictors. In the absence of official government records for 2020 for Sindhudurg district, our preliminary results using the CART algorithm compare reasonably well with 2018 government statistics on cashew extent. However, this is an ongoing study, and future analyses should also evaluate the performance of other classifiers such as the support vector machines (SVM)[11] algorithm and use accuracy assessments to compare algorithms.

This project will enable an evaluation of the most suitable methods that can be used to map cashew plantations in the Western Ghats, which could form the basis for future remote sensing studies that use GEE to map and detect cashew land use expansion. Mapping cashew monocultures using GEE can be a useful tool for conservation and land use planning. In the future, GEE could be used to monitor deforestation and cashew expansion and serve as a useful mapping tool for potential cashew sustainability certification programs.

---

7 See: http://www.fao.org/faostat/en/#data/QC (accessed 10 February, 2021).

8 According to Conservation International, an American non-profit environmental organization, to be qualified as a biodiversity hotspot, a region must have (a) at least 1,500 vascular plants as endemics, and (b) 30

9 RF is a non-parametric algorithm that uses a combination of decision trees to select the most suitable class for the pixel in question, while CART is a supervised algorithm that creates a binary decision tree and is more robust to outliers in predictors (Shaharum et al. 2020).

10 This software is now freely accessible. See: https://www.google.com/earth/versions/

11 SVM Constructs a 'hyperplane' or set of hyperplanes (hyperplane: a sub-space whose dimensions are one less than that of its model environment) in an infinite-dimensional space, which can further be used for image classification. See: https://scikit-learn.org/stable/modules/svm.html (accessed 24 March, 2021).

Results (implementation, impacts, constraints, enabling environment)

GEE implementation and dissemination

The GEE web application was originally created by Google engineers and mostly used by academics based in the USA for forest and vegetation land use and land cover studies (Kumar and Mutanga 2018). GEE was first developed to facilitate access to and analysis of large amounts of remote sensing data. However, the first users had to learn the programming language of the software and develop their own code and program to run land cover analyses. Academics were the principal users and developers of GEE since they had the required knowledge to create and implement a GEE program in their research. For example, the two-mapping projects described above were carried out by graduate students. The large effort of academics in developing GEE for land cover mapping and forest monitoring generated a new scientific literature about GEE implementation and thereby facilitated capacity building for individuals, who disseminated the technology to the private sector. Unfortunately, GEE has yet to be widely adopted by non-profit organizations (NGOs) or government agencies due to the knowledge gap in programming skills and geographical information system (GIS) software.

Potential contribution of GEE to sustainable forest management

GEE can have an impact by making satellite imagery accessible to a wide range of users and providing access to powerful computational tools for remote sensing through its web application. The use of remote sensing to generate information on land cover change, in particular on agricultural expansion into tropical forests, enhances our understanding of these processes. GEE can help monitor land cover changes and identify the main drivers of deforestation. However, on-the-ground interviews and field data collection is necessary to ground-truth the map resulting from remote sensing analysis using GEE and to gain a deeper understanding of causal relationships and responsible actors driving the process of deforestation. A combination of field-based data collection and tracking land cover change through satellite imagery could be used to inform future policies addressing land use changes and deforestation.

GEE can also enhance forest governance and enable forest owners or civil society organizations to detect changes in forest land cover and report these changes to government agencies. Civil society organizations involved in forest protection and climate change policies (e.g. FLEGT – Forest Law Enforcement, Governance and Trade, REDD – Reducing Emissions from Deforestation and Forest Degradation, NDC – Nationally Determined Contributions of the Paris Agreement) could hold governments and companies accountable for trends in deforestation in countries and corporate sectors respectively.

Constraints to dissemination and adoption

The main constraint to implementation is the limited number of people proficient in using GIS software. GEE itself is not complicated to implement once the code is accessible; however, users need basic knowledge of GIS to use the software. While using a GIS software, users can format the retrieved maps from GEE and integrate them into reports for policy advocacy. Countries and organizations who want to use maps and geographic information in policymaking and advocacy require people with such expertise. Launched in 2010, GEE is a comparatively young platform and is yet to gain widespread use and popularity. Future efforts can be focused on reaching out across all sections of academia and policymakers alike so that the multi-faceted benefits of GEE can be realized in practice. A more costly solution is to develop a third-party software based on GEE to guide users in implementing a specific code. A handful of software based on GEE for land monitoring already exist: for example, Collect Earth, which provides users with high-resolution satellite imagery to support the monitoring of agricultural lands, quantifying deforestation and land use changes (Saah et al. 2019).

Enabling factors to facilitate implementation

A handful of companies and organizations have already developed remote sensing tools to monitor forest evolution. GEE and other tools like Terra-I and Global Forest Watch[12] are already being used to monitor forest cover. In 22 African countries that have subscribed to warnings from services based on satellite imagery to detect decreases in forest cover, the probability of deforestation decreased by 18 percent in two years compared to the period from 2011 to 2016. Thanks to such warning systems, government agencies were better able to enforce existing forest protection laws or contribute to policies that reduced deforestation (Mofette et al. 2021).

Conclusions and wider implications of findings

People able to use GEE or equivalent remote sensing technologies can play a key role in influencing land use planning. For example, in Indonesia and India, such expertise has been used to map deforestation 'hotspots' and to identify the main drivers of this deforestation. Supported by the international community, civil society organizations can hire staff with GIS skills to analyze land use imagery to identify illegal deforestation and hold the guilty parties accountable.

GEE and remote sensing mapping tools can help future sustainability programs monitor oil palm and cashew land use expansion to better understand prior land use for commodity expansion, a key requirement when assessing corporate compliance with sustainability standards. It can hence be an effective landscape management tool for conservationists, regulators, and policymakers.

International donors should promote projects using such remote sensing tools for forest protection and advocacy for better environmental policies. Policy instruments like FLEGT, REDD+, and the Paris Agreement should seize the potential of GEE and remote sensing technologies to hold governments accountable for forest protection. For example, the Measurement, Reporting and Verification (MRV) provisions in the Paris Agreement should require land cover maps for mitigation based on land use.

12 See: https://www.globalforestwatch.org/

The views expressed in this information product are those of the author(s) and do not necessarily reflect the views or policies of FAO or CIFOR/FTA.

References

Comprehensive District Agriculture Plan of Sindhudurg. 2012. Department of Agriculture, Sindhudurg, Government of Maharashtra. http://krishi.maharashtra.gov.in/Site/Upload/Pdf/sindhudurg_cdap.pdf

Directorate of Cashewnut and Cocoa Development, Government of India. 2018. Available from: https://www.dccd.gov.in/Content.aspx?mid=1077&tid=1 (accessed 20 Juillet, 2021)

Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A, Thau D, Stehman SV, Goetz SJ, Loveland TR, et al. 2013. High-resolution global maps of 21st-century forest cover change. *Science* 342(6160): 850–853. https://doi.org/10.1126/science.1244693

Joshi N, Baumann M, Ehammer A, Fensholt R, Grogan K, Hostert P, Jepsen MR, Kuemmerle T, Meyfroidt P, Mitchard ETA, et al. 2016. A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. *Remote Sensing* 8(1): 70. https://doi.org/10.3390/rs8010070

Kumar L and Mutanga O. 2018. Google Earth Engine applications since inception: Usage, trends, and potential. *Remote Sensing* 10(10): 1509. https://doi.org/10.3390/rs10101509

Moffette F, Alix-Garcia J, Shea K and Pickens AH. 2021. The impact of near-real-time deforestation alerts across the tropics. *Nature Climate Change* 11: 172–178. https://doi.org/10.1038/s41558-020-00956-w

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB and Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858. https://doi.org/10.1038/35002501

Punjabi GA and Kulkarni J. 2015. Examining large carnivore connectivity and creating conservation networks in the Sahyadri-Konkan corridor. Final report. Pune, India: Wildlife Research & Conservation Society and The Nityata Foundation. https://indiabiodiversity.org/biodiv/content/projects/project-489b617d-e927-42d0-ad57-63e90f636b83/13.pdf

Saah D, Johnson G, Ashmall B, Tondapu G, Tenneson K, Patterson M, Poortinga A, Markert K, Nguyen HQ, Aung KS, et al. 2019. Collect Earth: An online tool for systematic reference data collection in land cover and use applications. *Environmental Modelling & Software* 118: 166–171. https://doi.org/10.1016/j.envsoft.2019.05.004

Sarzynski T, Giam X, Carrasco L and Lee JSH. 2020. Combining radar and optical imagery to map oil palm plantations in Sumatra, Indonesia, using the Google Earth Engine. *Remote Sensing* 12(7): 1220. https://doi.org/10.3390/rs12071220

Sengar SH, Mohod AG and Khandetod YP. 2012. Performance evaluation of kiln for cashew nut shell carbonization and liquid. *International Journal of Energy Engineering* 2(3): 78–85. https://doi.org/10.5923/j.ijee.20120203.04

Shaharum NSN, Shafri HZM, Ghani WAWAK, Samsatli S, Al-Habshi MMA and Yusuf B. 2020. Oil palm mapping over Peninsular Malaysia using Google Earth Engine and machine learning algorithms. *Remote Sensing Applications: Society and Environment* 17: 100287. https://doi.org/10.1016/j.rsase.2020.100287